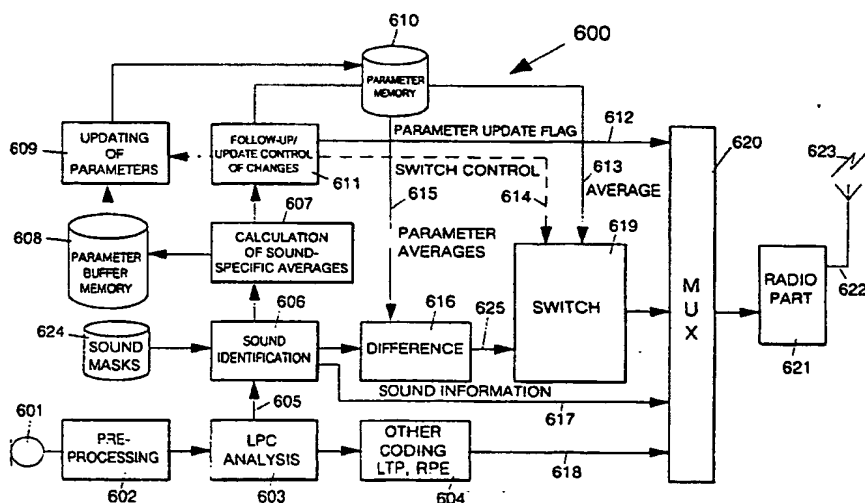


## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>5</sup> : <b>G10L 9/02</b>	<b>A1</b>	(11) International Publication Number: <b>WO 94/18668</b> (43) International Publication Date: <b>18 August 1994 (18.08.94)</b>
<p>(21) International Application Number: <b>PCT/FI94/00051</b></p> <p>(22) International Filing Date: <b>3 February 1994 (03.02.94)</b></p> <p>(30) Priority Data: <b>930493</b>      <b>4 February 1993 (04.02.93)</b>      <b>FI</b></p> <p>(71) Applicant (for all designated States except US): <b>NOKIA TELECOMMUNICATIONS OY [FI/FI]; Mäkkylän puistotie 1, FIN-02600 Espoo (FI).</b></p> <p>(72) Inventor; and (75) Inventor/Applicant (for US only): <b>VÄNSKÄ, Marko [FI/FI]; Päistöspolku B 10, FIN-03100 Nummela (FI).</b></p> <p>(74) Agent: <b>OY KOLSTER AB; Iso Roobertinkatu 23, P.O. Box 148, FIN-00121 Helsinki (FI).</b></p>		<p>(81) Designated States: <b>AU, CN, GB, JP, NO, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).</b></p> <p><b>Published</b> <i>With international search report.</i></p>

(54) Title: A METHOD OF TRANSMITTING AND RECEIVING CODED SPEECH



## (57) Abstract

A method of transmitting (600) and receiving coded speech, in which method samples are taken (602) of a speech signal (601) and reflection coefficients are calculated (603) from these samples. In order to minimize the used transmission rate, characteristics of the reflection coefficients are compared (606) with respective stored (624) sound-specific characteristics of the reflection coefficients for the identification of the sounds, and identifiers of identified sounds are transmitted (617), speaker-specific characteristics are calculated (607) for the reflection coefficients representing the same sound and stored in a memory (608, 609, 610), the calculated characteristics of the reflection coefficients representing said sound and stored in the memory (610) are compared (611) with the following characteristics of the reflection coefficients representing the same sound, and if the following characteristics of the reflection coefficients representing the same sound do not essentially differ from the characteristics of the reflection coefficients stored in the memory (610), differences between the characteristics of the reflection coefficients representing the same sound of the speaker and the characteristics of the reflection coefficients calculated from the previous sample are calculated (616) and transmitted (625).

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

## A method of transmitting and receiving coded speech

### Field of the Invention

5 The invention relates to a method of transmitting coded speech, in which method samples are taken of a speech signal and reflection coefficients are calculated from these samples.

The invention relates also to a method of receiving coded speech.

10

### Background of the Invention

In telecommunication systems, especially on the radio path of radio telephone systems, such as GSM system, it is known that a speech signal entering the system and to be transmitted is preprocessed, i.e. filtered and converted into digital form. In known systems the signal is then coded by a suitable coding method, e.g. by the LTP (Long Term Prediction) or RPE (Regular Pulse Excitation) method. The GSM system typically uses a combination of these, i.e. the RPE-LTP method, which is described in detail e.g. in "M. Mouly and M.B. Paute, The GSM System for Mobile Communications, 1992, 49, rue PALAISEAU F-91120, pages 155 to 162". These methods are described in more detail in the GSM Specification "GSM 06.10, January 1990, GSM Full Rate Speech Transcoding, ETSI, 93 pages".

25 A drawback of the known techniques is the fact that the coding methods used require plenty of transmission capacity. When using these methods according to the prior art, the speech signal to be transmitted to the receiver has to be transmitted entirely, whereby transmission capacity is unnecessarily wasted.

35

## Disclosure of the Invention

The object of this invention is to offer such a speech coding method for transmitting data in telecommunication systems by which the transmission speed required for speech transmission may be lowered and/or the required transmission capacity may be reduced.

This novel method of transmitting coded speech is provided by means of the method of the invention, which is characterized in that characteristics of the reflection coefficients are compared with respective sound-specific characteristics of the reflection coefficients of at least one previous speaker for the identification of the sounds and identifiers of the identified sounds are transmitted, speaker-specific characteristics are calculated for the reflection coefficients representing the same sound and stored in a memory, the calculated characteristics of the reflection coefficients representing the same sound and stored in the memory are compared with the following characteristics of the reflection coefficients representing the same sound, and if the following characteristics of the reflection coefficients representing the same sound differ essentially from the characteristics of the reflection coefficients stored in the memory, the new characteristics representing the same sound are stored in the memory and transmitted, and before transmitting them, an information is sent of the transmission of these characteristics and if the following characteristics of the reflection coefficients representing the same sound do not essentially differ from the characteristics of the reflection coefficients stored in the memory, differences between the characteristics of the reflection coefficients representing the same sound of the speaker and the characteristics of the reflection coefficients

calculated from the previous sample are calculated and transmitted.

5 The invention relates further to a method of receiving coded speech, which method is characterized in that an identifier of an identified sound is received, differences between characteristics of the stored sound-specific reflection coefficients of one previous speaker and characteristics of the reflection coefficients calculated from samples are re-  
10 ceived, the speaker-specific characteristics of the reflection coefficients corresponding to the received sound identifier are searched for in a memory and added to said differences, and from this sum are calculated new reflection coefficients used for sound  
15 production, and if an information of a transmission of new characteristics sent by a communications transmitter as well as new characteristics of the reflection coefficients representing the same sound sent by another communications transmitter are re-  
20 ceived, these new characteristics are stored in the memory.

The invention is based on the idea that, for a transmission, a speech signal is analyzed by means of the LPC (Linear Prediction Coding) method, and a set  
25 of parameters, typically characteristics of reflection coefficients, modelling a speaker's vocal tract is created for the speech signal to be transmitted. According to the invention, sounds are then identified from the speech to be transmitted by comparing  
30 the reflection coefficients of the speech to be transmitted with several speakers' respective previously received reflection coefficients calculated for the same sound. After this, reflection coefficients and some characteristics therefor are calculated for  
35 each sound of the speaker concerned. Characteristic

may be a number representing physical dimensions of a lossless tube modelling the speaker's vocal tract. Subsequently, from these characteristics are subtracted the characteristics of the reflection coefficients corresponding to each sound, providing a difference, which is transmitted to the receiver together with an identifier of the sound. Before that, information of the characteristics of the reflection coefficients corresponding to each sound identifier has been transmitted to the receiver, and therefore, the original sound may be reproduced by summing said difference and the previously received characteristic of the reflection coefficients, and thus, the amount of information on the transmission path decreases.

Such a method of transmitting and receiving coded speech has the advantage that less transmission capacity is needed on the transmission path, because each speaker's all voice properties need not be transmitted, but it is enough to transmit the identifier of each sound of the speaker and the deviation by which each separate sound of the speaker deviates from a property, typically an average, of some characteristic of the previous reflection coefficients of each sound of this speaker. By means of the invention, it is thus possible to reduce the transmission capacity needed for speech transmission by approximately 10 % in total, which is a considerable amount.

In addition, the invention may be used for recognizing the speaker in such a way that some characteristic, for instance an average, of the speaker's sound-specific reflection coefficients is stored in a memory in advance, and the speaker is then recognized, if desired, by comparing the characteristics of the reflection coefficients of some sound of the speaker with said characteristic calculated in ad-

vance.

Cross-sectional areas of cylinder portions of a lossless tube model used in the invention may be calculated easily from so-called reflection coefficients produced in conventional speech coding algorithms. Also some other cross-sectional dimension, such as radius or diameter, may naturally be determined from the area to constitute a reference parameter. On the other hand, instead of being circular the cross-section of the tube may also have some other shape.

#### Description of the Drawings

In the following, the invention will be described in more detail with reference to the attached drawings, in which

Figures 1 and 2 illustrate a model of a speaker's vocal tract by means of a lossless tube comprising successive cylinder portions,

Figure 3 illustrates how the lossless tube models change during speech, and

Figure 4 shows a flow chart illustrating identification of sounds,

Figure 5a is a block diagram illustrating speech coding on a sound level in a transmitter according to the invention,

Figure 5b shows a transaction diagram illustrating a reproduction of a speech signal on a sound level in a receiver according to the invention,

Figure 6 shows a communications transmitter implementing the method according to the invention,

Figure 7 shows a communications receiver implementing the method according to the invention.

#### Detailed Description of the Invention

Reference is now made to Figure 1 showing a

perspective view of a lossless tube model comprising successive cylinder portions C1 to C8 and constituting a rough model of a human vocal tract. The lossless tube model of Figure 1 can be seen in side view in Figure 2. The human vocal tract generally refers to a vocal passage defined by the human vocal cords, the larynx, the mouth of pharynx and the lips, by means of which tract a man produces speech sounds. In the Figures 1 and 2, the cylinder portion C1 illustrates the shape of a vocal tract portion immediately after the glottis between the vocal cords, the cylinder portion C8 illustrates the shape of the vocal tract at the lips and the cylinder portions C2 to C7 inbetween illustrate the shape of the discrete vocal tract portions between the glottis and the lips. The shape of the vocal tract typically varies continuously during speaking, when sounds of different kinds are produced. Similarly, the diameters and areas of the discrete cylinders C1 to C8 representing the various parts of the vocal tract also vary during speaking. However, a previous patent application FI-912088 of this same inventor discloses that the average shape of the vocal tract calculated from a relatively high number of instantaneous vocal tract shapes is a constant characteristic of each speaker, which constant may be used for a more compact transmission of sounds in a telecommunication system or for recognizing the speaker. Correspondingly, the averages of the cross-sectional areas of the cylinder portions C1 to C8 calculated in the long term from the instantaneous values of the cross-sectional areas of the cylinders C1 to C8 of the lossless tube model of the vocal tract are also relatively exact constants. Furthermore, the values of the cross-sectional dimensions of the cylinders are also determined by the values of



the actual vocal tract and are thus relatively exact constants characteristic of the speaker.

The method according to the invention utilizes so-called reflection coefficients produced as a provisional result at Linear Predictive Coding (LPC) well-known in the art, i.e. so-called PARCOR-coefficients  $r_k$  having a certain connection with the shape and structure of the vocal tract. The connection between the reflection coefficients  $r_k$  and the areas  $A_k$  of the cylinder portions  $C_k$  of the lossless tube model of the vocal tract is according to the formula (1)

$$r(k) = \frac{A(k+1) - A(k)}{A(k+1) + A(k)} \quad (1)$$

where  $k = 1, 2, 3, \dots$ . Such a cross-sectional area can be considered as a characteristic of a reflection coefficient.

The LPC analysis producing the reflection coefficients used in the invention is utilized in many known speech coding methods. One advantageous embodiment of the method according to the invention is expected to be coding of speech signals sent by subscribers in radio telephone systems, especially in the Pan-European digital radio telephone system GSM. The GSM Specification 06.10 defines very accurately the LPC-LTP-RPE (Linear Predictive Coding - Long Term Prediction - Regular Pulse Excitation) speech coding method used in the system. It is advantageous to use the method according to the invention in connection with this speech coding method, because the reflection coefficients needed in the invention are obtained as a provisional result from the above-mentioned prior art LPC-RPE-LTP coding method. In the

invention, the steps of the method follow said speech coding algorithm complying with the GSM Specification 06.10 up to the calculation of the reflection coefficients, and as far as the details of these steps are concerned, reference is made to said specification. In the following, these method steps will be described only generally in those parts which are essential for the understanding of the invention with reference to the flow chart of Figure 4.

10 In Figure 4, an input signal IN is sampled in block 10 at a sampling frequency 8 kHz, and an 8-bit sample sequence  $s_i$  is formed. In block 11, a DC component is extracted from the samples so as to eliminate an interfering side tone possibly occurring in coding. After this, the sample signal is pre-emphasized in block 12 by weighting high signal frequencies by a first-order FIR (Finite Impulse Response) filter. In block 13 the samples are segmented into frames of 160 samples, the duration of each frame being about 20 ms.

20 In block 14, the spectrum of the speech signal is modelled by performing an LPC analysis on each frame by an auto-correlation method, the performance level being  $p=8$ .  $p+1$  values of the auto-correlation function ACF are then calculated from the frame by means of the formula (2) as follows:

$$30 \quad \text{ACF}(k) = \sum_{i=1}^{160} s(i)s(i-k) \quad (2)$$

where  $k = 0, 1, \dots, 8$ .

Instead of the auto-correlation function, it is possible to use some other suitable function, such as a co-variance function. The values of eight so-called

reflection coefficients  $r_k$  of a short-term analysis filter used in a speech coder are calculated from the obtained values of the auto-correlation function by Schur's recursion 15 or some other suitable recursion method. Schur's recursion produces new reflection coefficients every 20th ms. In one embodiment of the invention the coefficients comprise 16 bits and their number is 8. By applying Schur's recursion 15 for a longer time, the number of the reflection coefficients can be increased, if desired.

In step 16, a cross-sectional area  $A_k$  of each cylinder portion  $C_k$  of the lossless tube modelling the speaker's vocal tract by means of the cylindrical portions is calculated from the reflection coefficients  $r_k$  calculated from each frame. As Schur's recursion 15 produces new reflection coefficients every 20th ms, 50 cross-sectional areas per second will be obtained for each cylinder portion  $C_k$ . After the cross-sectional areas of the cylinders of the lossless tube have been calculated, the sound of the speech signal is identified in step 17 by comparing these calculated cross-sectional areas of the cylinders with the values of the cross-sectional areas of the cylinders stored in a parameter memory. This comparing operation will be presented in more detail in connection with the explanation of Figure 5 referring to reference numerals 60, 60A and 61, 61A. In step 18, average values  $A_{k,ave}$  of the areas of the cylinder portions  $C_k$  of the lossless tube model are calculated for a sample taken of the speech signal, and the maximum cross-sectional area  $A_{k,max}$  occurred during the frames is determined for each cylinder portion  $C_k$ . Then in step 19, the calculated averages are stored in a memory, e.g. in a buffer memory 608 for parameters, shown below in Figure 6. Subsequently, the

averages stored in the buffer memory 608 are compared with the cross-sectional areas of the just obtained speech samples, in which comparison is calculated whether the obtained samples differ too much from the previously stored averages. If the obtained samples differ too much from the previously stored averages, an updating 21 of the parameters, i.e. the averages, is performed, which means that a follow-up and update block 611 of changes controls a parameter update block 609 in the way shown in Figure 6 to read the parameters from the parameter buffer memory 608 and to store them in a parameter memory 610. Simultaneously, those parameters are transmitted via a switch 619 to a receiver, the structure of which is illustrated in Figure 7. On the other hand, if the obtained samples do not differ too much from the previously stored averages, the parameters of an instantaneous speech sound obtained from the sound identification shown in Figure 6 are supplied to a subtraction means 616. This takes place in step 22 of Figure 4, in which the subtraction means 616 searches in the parameter memory 610 for the averages of the previous parameters representing the same sound and subtracts from them the instantaneous parameters of the just obtained sample, thus producing a difference, which is transmitted 625 to the switch 619 controlled by the follow-up and update block 611 of changes, which switch sends forward the difference signal via a multiplexer 620 MUX to the receiver in step 23. This transmission will be described more accurately in connection with the explanation of Figure 6. The follow-up and update block 611 of changes controls the switch 619 to connect the different input signals, i.e. the updating parameters or the difference, to the multiplexer 620 and a radio part 621 in a way

appropriate in each case.

In the embodiment of the invention shown in Figure 5a, the analysis used for speech coding on a sound level is described in such a way that the averages of the cross-sectional areas of the cylinder portions of the lossless tube modelling the vocal tract are calculated from a speech signal to be analyzed, from the areas of the cylinder portions of instantaneous lossless tube models created during a predetermined sound. The duration of one sound is rather long, so that several, even tens of temporally consecutive lossless tube models can be calculated from a single sound present in the speech signal. This is illustrated in Figure 3, which shows four temporally consecutive instantaneous lossless tube models S1 to S4. From Figure 3 can be seen clearly that the radii and cross-sectional areas of the individual cylinders of the lossless tube vary in time. For instance, the instantaneous models S1, S2 and S3 could roughly classified be created during the same sound, so that their average could be calculated. The model S4, instead, is clearly different and associated with another sound and therefore not taken into account in the averaging.

In the following, speech coding on a sound level will be described with reference to the block diagram of Figure 5a. Even though speech coding can be made by means of a single sound, it is reasonable to use in the coding all those sounds the communicating parties wish to send to each other. All vowels and consonants can be used, for instance.

The instantaneous lossless tube model 59 created from a speech signal can be identified in block 52 to correspond to a certain sound, if the cross-sectional dimension of each cylinder portion of the in-

stantaneous lossless tube model 59 is within the pre-determined stored limit values of the corresponding sound of a known speaker. These sound-specific and cylinder-specific limit values are stored in a so-called quantization table 54 creating a so-called sound mask included in a memory means indicated by the reference numeral 624 in Figure 6. In Figure 5a, the reference numerals 60 and 61 illustrate how said sound- and cylinder-specific limit values create a mask or model for each sound, within the allowed area 60A and 61A (unshaded areas) of which the instantaneous vocal tract model 59 to be identified has to fit. In Figure 5a, the instantaneous vocal tract model 59 fits the sound mask 60, but does obviously not fit the sound mask 61. Block 52 thus acts as a kind of sound filter, which classifies the vocal tract models into correct sound groups a, e, i, etc. After the sounds have been identified in block 606 of Figure 6, i.e. in step 52 of Figure 5a, the parameters corresponding to the identified sounds a, e, i, k are stored in the buffer memory 608 of Figure 6, to which memory corresponds block 53 of Figure 5a. From this buffer memory 608, or block 53 of Figure 5a, the sound parameters are stored further under the control of the follow-up and update control block of changes of Figure 6 in an actual parameter memory 55, in which each sound, such as a, e, i, k, has parameters corresponding to that sound. At the identification of sounds, it has also been possible to provide each sound to be identified with an identifier, by means of which the parameters corresponding to each instantaneous sound can be searched for in the parameter memory 55, 610. These parameters can be supplied to the subtraction means 616, which calculates 56 according to Figure 5a the difference be-

tween the parameters of the sound searched for in the parameter memory by means of the sound identifier and the instantaneous values of this sound. This difference will be sent further to the receiver in the manner shown in Figure 6, which will be described in more detail in connection with the explanation of that figure.

Figure 5b is a transaction diagram illustrating a reproduction of a speech signal on a sound level according to the invention, taking place in a receiver. The receiver receives an identifier 500 of a sound identified by a sound identification unit (reference numeral 606 in Figure 6) of the transmitter and searches in its own parameter memory 501 (reference numeral 711 in Figure 7), on the basis of the sound identifier 500, for the parameters corresponding to the sound and supplies 502 them to a summer 503 (reference numeral 712 in Figure 7) creating new characteristics of reflection coefficients by summing the difference and the parameters. By means of these numbers are calculated new reflection coefficients, from which can be calculated a new speech signal. Such a creation of speech signal by summing will be described in greater detail in Figure 7 and in the explanation attached to it.

Figure 6 shows a communications transmitter 600 implementing the method of the invention. A speech signal to be transmitted is supplied to the system via a microphone 601, from which the signal converted into electrical form is transmitted to a preprocessing unit 602, in which the signal is filtered and converted into digital form. Then an LPC analysis of the digitized signal is performed in an LPC analyzer 603, typically in a signal processor. The LPC analysis results in reflection coefficients 605, which are

led to the transmitter according to the invention. The rest of the information passed through the LPC analyzer is supplied to other signal processing units 604, performing the other necessary codings, such as LTP and RPE codings. The reflection coefficients 605 are supplied to a sound identification unit 606 comparing the instantaneous cross-sectional values of the vocal tract of the speaker creating the sound in question, which values are obtained from the reflection coefficients of the supplied sound, or other suitable values, an example of which is indicated by the reference numeral 59 in Figure 5, with the sound masks of the available sounds stored already earlier in a memory means 624. These masks are illustrated by the reference numerals 60, 60A, 61 and 61A in Figure 5. After the sounds uttered by the speaker have been successfully discovered from the information 605 supplied to the sound identification unit 606, averages corresponding to each sound are calculated for this particular speaker in a sound-specific averaging unit 607. The sound-specific averages of the cross-sectional values of the vocal tract of that speaker are stored in a parameter buffer memory 608, from which a parameter update block 609 stores the average of each new sound in a parameter memory 610 at updating of parameters. After the calculation of the sound-specific averages, the values corresponding to each sound to be analyzed, i.e. the values from the temporally unbroken series of which the average was calculated, are supplied to a follow-up and update control block 611 of changes. That block compares the average values of each sound stored in the parameter memory 610 with the previous values of the same sound. If the values of a just arrived previous sound differ sufficiently from the averages of the previous



sounds, an updating of the parameters, i.e. averages, is at first performed in the parameter memory, but these parameters, being the averages of the cross-sections of the vocal tract needed for the production of each sound, i.e. the averages 613 of the parameters, are also sent via a switch 619 to a multiplexer 620 and from there via a radio part 621 and an antenna 622 to a radio path 623 and further to a receiver. In order to inform the receiver of the fact that the information sent by the transmitter consists of updating information of parameters, the follow-up and update control block 611 of changes sends to the multiplexer 620 a parameter update flag 612, which is transmitted further to the receiver along the route 621, 622, 623 described above.

The switch 619 is controlled 614 by the follow-up and update control block 611 in such a way that the parameters pass through the switch 619 further to the receiver, when they are updated.

When new parameters have been sent to the receiver in a situation in which the communication has started, meaning that no parameters have been sent to the receiver earlier, or when new parameters replacing the old parameters have been sent to the receiver, a transmission of coded sounds begins at the arrival of next sound. The parameters of the sound identified in the sound identification unit 606 are then transmitted to the subtraction means 616. Simultaneously, an information of the sound 617 is transmitted via the multiplexer 620, the radio part 621, the antenna 622 and the radio path 623 to the receiver. This sound information may be for instance a bit string representing a fixed binary number. In the subtraction means 616, the parameters of the just identified 606 sound are subtracted from the aver-

ages 615 of the previous parameters representing the same sound, which averages have been searched for in the parameter memory 610, and the calculated difference is transmitted 625 via the multiplexer 620 along the route 621, 622, 623 described above further to the receiver. An attentive reader observes that the advantage obtained by the method of the invention, i.e. a reduction in the needed transmission capacity, is based on this very difference produced by subtraction and on the transmission of this difference.

Figure 7 shows a communications receiver 700 implementing the method of the invention. A signal transmitted by the communications transmitter 600 of Figure 6 via a radio path 623 = 701 or some other medium is received by an antenna 702, from which the signal is led to a radio part 703. If the signal sent by the transmitter 600 is coded in another way than by LPC coding, it is received by a demultiplexer 704 and transmitted to a means 705 for other decoding, i.e. LTP and RPE decoding. The sound information sent by the transmitter 600 is received by the demultiplexer 704 and transmitted 706 to a sound parameters searching unit 718. The information of updated parameters is also received by the demultiplexer 704 DEMUX and led to a switch 707 controlled by a parameter update flag 709 received in the same way. A subtraction signal sent by the transmitter 600 is also applied to the switch 707. The switch 707 transmits 710 the information of updated parameters, i.e. the new parameters corresponding to the sounds, to a parameter memory 711. The received difference between the averages of the sound just arrived and the previous parameters representing the same sound is transmitted 708 to a summer 712. The sound identifier, i.e. the sound information, was thus transmitted to

the sound parameters searching unit 718 searching 716 for the parameters corresponding to (the identifier of) the sound stored in the parameter memory 711, which parameters are transmitted 717 by the parameter  
5 memory 711 to the summer 712 for the calculation of the coefficients. The summer 712 sums the difference 708 and the parameters obtained 717 from the parameter memory 711 and calculates from them new coefficients, i.e. new reflection coefficients. By means of  
10 these coefficients is created a model of the vocal tract of the original speaker and speech is thus produced resembling the speech of this original speaker. The new calculated reflection coefficients are transmitted 713 to an LPC decoder 714 and further to a  
15 postprocessing unit 715 performing a digital/analog conversion and applying the amplified speech signal further to a loudspeaker 720, which reproduces the speech corresponding to the speech of the original speaker.

20 The above method according to the invention can be implemented in practice for instance by means of software, by utilizing a conventional signal processor.

25 The drawings and the explanation associated with them are only intended to illustrate the idea of the invention. As to the details, the method of the invention of transmitting and receiving coded speech may vary within the scope of the claims. Though the invention has above been described primarily in connection with radio telephone systems, especially the  
30 GSM mobile phone system, the method of the invention can be utilized also in telecommunication systems of other kinds.

## Claims:

1. A method of transmitting (600) coded speech,  
in which method samples are taken (10; 602) of a  
5 speech signal (IN; 601) and reflection coefficients  
are calculated (603) from these samples, the method  
being characterized in that

characteristics of the reflection coefficients  
are compared (17; 606) with respective stored (624;  
10 54) sound-specific characteristics of the reflection  
coefficients of at least one previous speaker for the  
identification of the sounds, and identifiers of the  
identified sounds are transmitted (617),

speaker-specific characteristics are calculated  
15 (607) for the reflection coefficients representing  
the same sound and stored in a memory (608, 609,  
610),

the calculated characteristics of the re-  
flection coefficients representing said sound and  
20 stored in the memory (610) are compared (20; 611)  
with the following characteristics of the reflection  
coefficients representing the same sound, and if the  
following characteristics of the reflection coeffi-  
cients representing the same sound differ (21) essen-  
25 tially from the characteristics of the reflection co-  
efficients stored in the memory (610), the new char-  
acteristics representing the same sound are stored  
(609) in the memory (610) and transmitted (613), and  
before transmitting them, an information (612) is  
30 sent of the transmission of these characteristics,

and if the following characteristics of the  
reflection coefficients representing the same sound  
do not essentially differ (20) from the characteris-  
tics of the reflection coefficients stored in the  
35 memory (610), differences between the characteristics

of the reflection coefficients representing the same sound of the speaker and the characteristics of the reflection coefficients calculated from the previous sample are calculated and transmitted (625).

5           2. A method of receiving (700) coded speech, which method is c h a r a c t e r i z e d in that an identifier of an identified sound is received (706; 500),

10           differences (708) between characteristics of the stored sound-specific reflection coefficients of one previous speaker and characteristics of the reflection coefficients calculated from samples are received,

15           the speaker-specific characteristics of the reflection coefficients corresponding to the received sound identifier are searched for (718, 716) in a memory (711; 501) and added (712; 503) to said differences (708), and from this sum are calculated new reflection coefficients (713) used for sound (720) production, and

20           if an information (709) of a transmission of new characteristics sent by a communications transmitter (600) as well as new characteristics (710) of the reflection coefficients representing the same sound sent by another communications transmitter are received, these new characteristics are stored in the memory (711).

25           3. A method according to claim 1 or 2, c h a r a c t e r i z e d in that said characteristics are averages of the reflection coefficients.

30

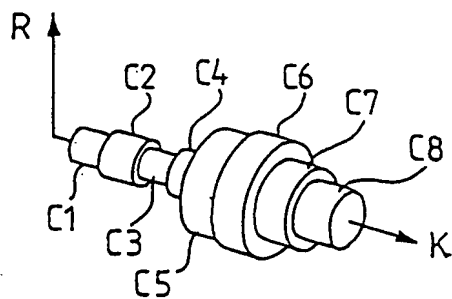


FIG. 1

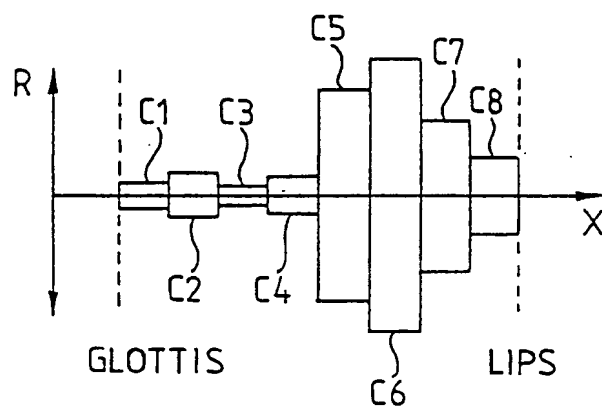


FIG. 2

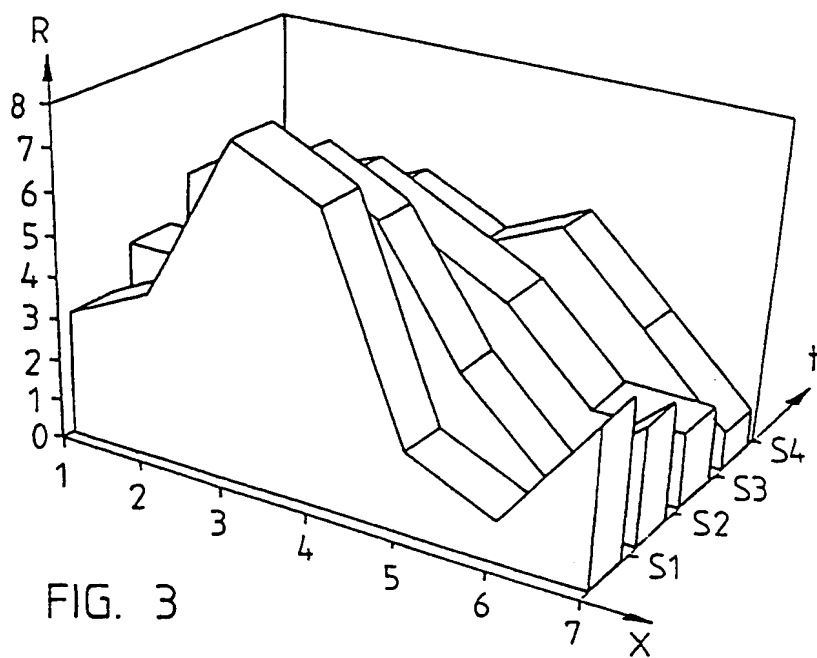


FIG. 3

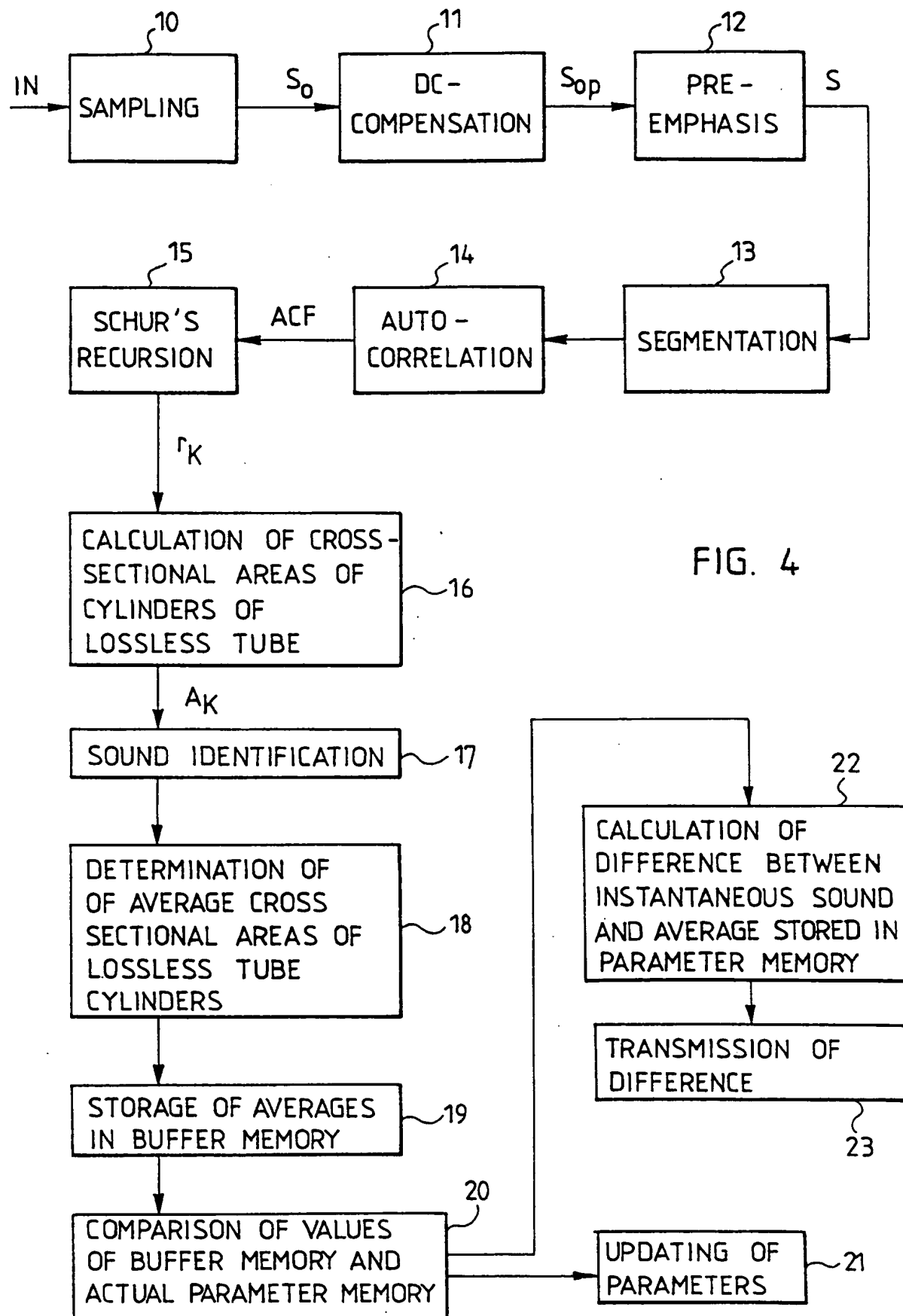


FIG. 4

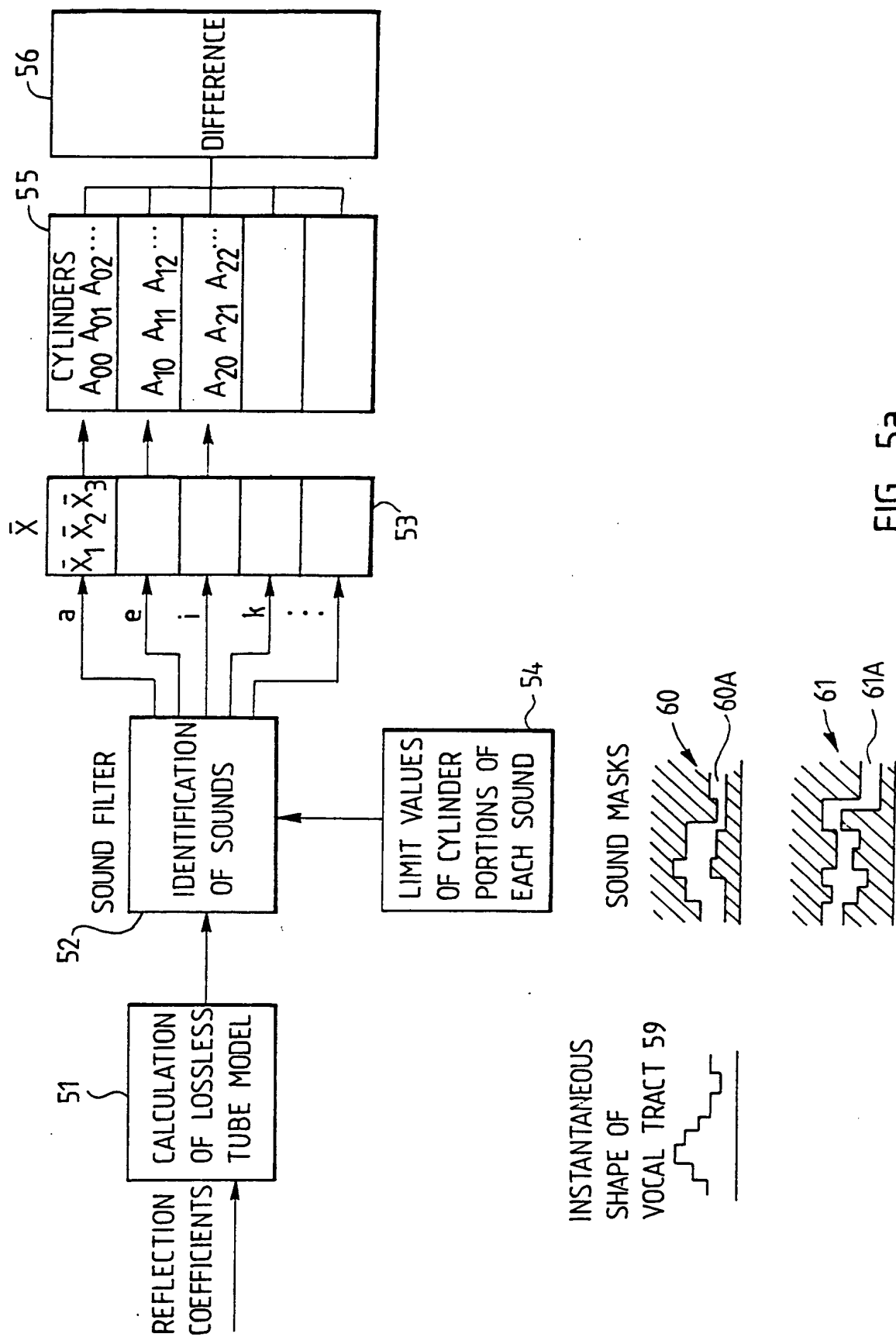


FIG. 5a



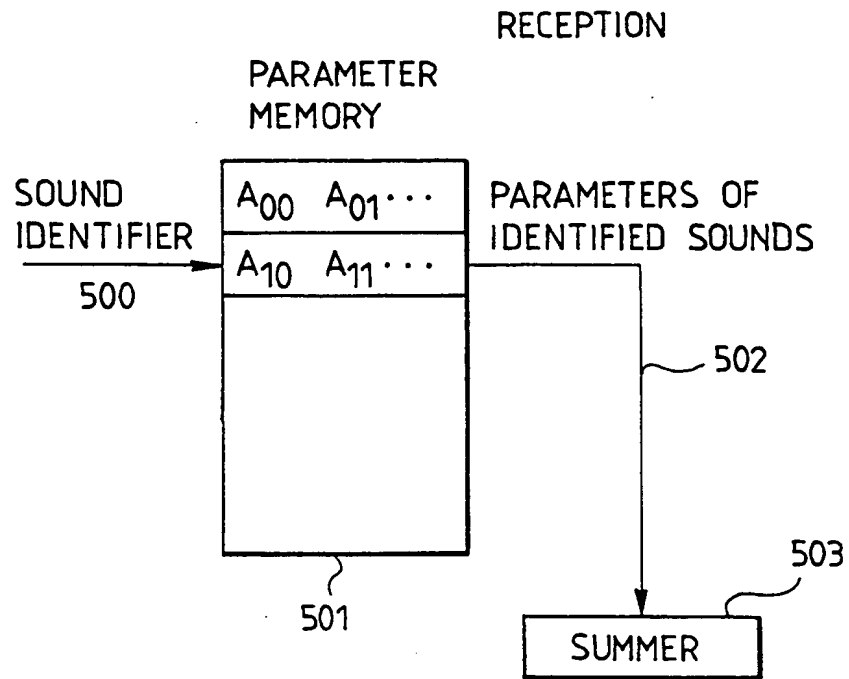
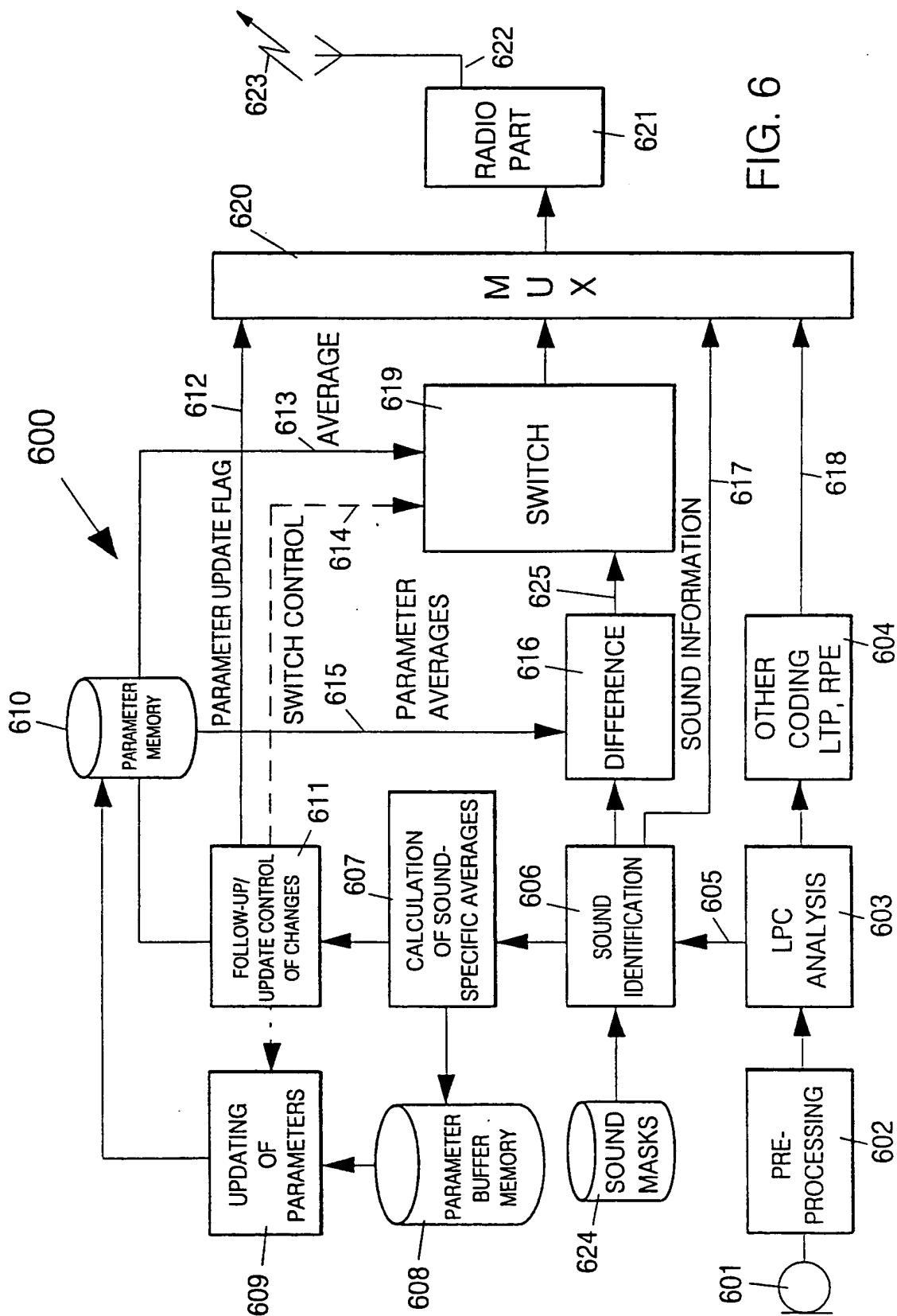
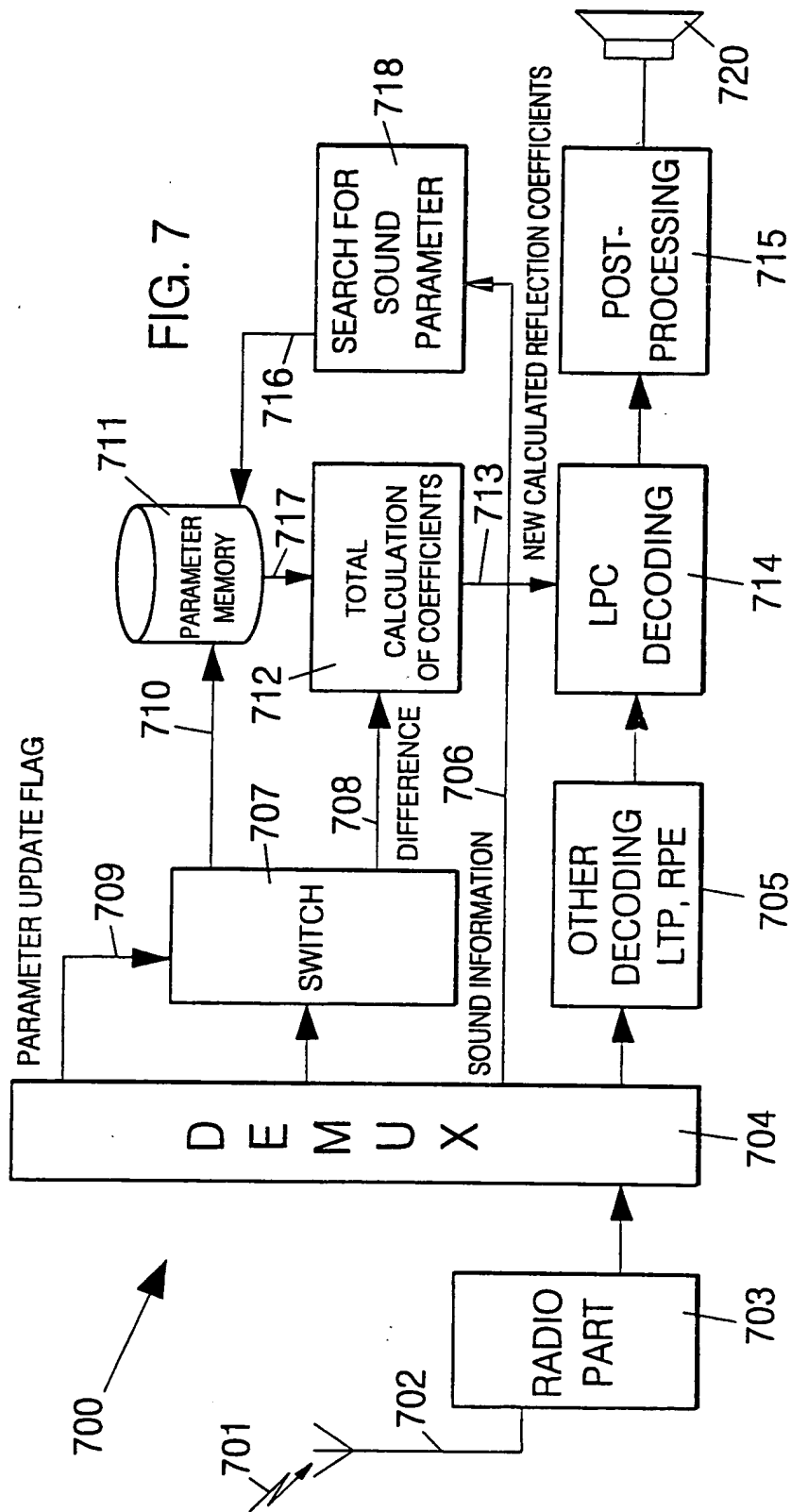


FIG. 5b





## INTERNATIONAL SEARCH REPORT

International application No.

PCT/FI 94/00051

## A. CLASSIFICATION OF SUBJECT MATTER

IPC <sup>6</sup>: G10L 9/02

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC <sup>6</sup>: G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPIL, EDOC

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US, A, 5121434 (CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE), 9 June 1992 (09.06.92), column 1, line 64 - column 2, line 7; column 9, line 4 - line 18, abstract --	1,2
A	WO, A1, 9220064 (TELENOKIA OY), 12 November 1992 (12.11.92), see the whole document --	1-3
A,P	WO, A1, 9402936 (VOICE POWERED TECHNOLOGY INTERNATIONAL INC.), 3 February 1994 (03.02.94), page 4, line 5 - page 6, line 26, claims 1-8 -- -----	1,2

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

\* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&amp;" document member of the same patent family

Date of the actual completion of the international search

3 May 1994

Date of mailing of the international search report

19 -05- 1994

Name and mailing address of the ISA/  
Swedish Patent Office  
Box 5055, S-102 42 STOCKHOLM  
Facsimile No. +46 8 666 02 86

Authorized officer

RUNE BENGTSOON

Telephone No. +46 8 782 25 00

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

16/04/94

International application No.

PCT/FI 94/00051

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
US-A-	5121434	09/06/92	EP-A-	0347338	20/12/89
WO-A1-	9220064	12/11/92	AU-A-	1653092	21/12/92
			EP-A-	0537316	21/04/93
WO-A1-	9402936	03/02/94	NONE		

This Page Blank (uspto)